Spring 2020

The Role of Socio-Economic Status on the Language Output of 3-6 year old's during a Naturalistic Museum Study

Serena K. Mueller
University of New Hampshire - Main Campus

The Role of Socio-Economic Status on the Language Output of 3-6-year old's during a Naturalistic Museum Study

Serena Mueller

The University of New Hampshire
105 Main St.
Durham, NH 03824

Abstract

Over the past 50 years, the lexical nature of young children's multi-word utterances has been disputed. We ask how socioeconomic status (SES) impacts content and function word usage, and how function word use impacts MLU. Findings show that SES positively correlates with content word production and MLU, which contributes to our knowledge on how SES shapes language and highlights the need for its consideration across settings.

Introduction

Children learn language at various rates but tend to conform to the same general timeline in terms of when they produce their first word, when they begin to combine words, as well as the types of grammatical markers they use in their speech (Benedict, 1979; Brown & Berko, 1960; Fenson at al., 1994). However, the varied rates at which individuals develop are based on differences in their innate capacities and their environments. SES affects the rate at which children develop language and can be defined as a complex measurement including individual factors such as social status, income, education, and occupation or a combination of these elements (Duncan & Magnuson, 2013; Hackman & Farah, 2009). Previous research has suggested that SES can affect the quantity as well as quality of adult speech input that a child receives (Hart & Risley, 1995). Children best learn language by being exposed to it, and they will learn a greater variety of words if they are exposed to not only more speech, but high-quality caliber language use in different settings. This may have an impact on the types of words that they use in their speech most frequently, and by hearing language in various settings, children can generalize new words and their meanings to use in novel contexts (Hart et al., 1995).

There is emerging evidence of how SES impacts not only language development, but how it influences a child's production and use of language. Little is known about the specific

SES-related variation in children's language production as few studies have assessed the differences in the types of words produced by children using American English. To explore these connections, Le Normand et al. (2013) investigated how SES influenced the development and use of grammar and mean length of utterance (MLU) in French-acquiring children. They found SES did not have a significant impact on the grammatical word types produced across the age groups but suggests that similar studies should reduplicate this study in other languages to evaluate the generalizability of the findings.

This current study tests the findings of Le Normand et al. (2013)'s as it is the first time that native American English-speaking children from the seacoast region of New Hampshire and Maine are analyzed from naturalistic language samples to help generalize the affects of SES on word production. This paper presents a comprehensive study of the impact SES has on word-type production, specifically content and function words. Critically, the central aim is to explore the language development of children being raised in the seacoast area of New Hampshire and Maine and to analyze how SES affects the early processes of language development such as specific word use, as well as the ability to develop vocabulary based on the amount and type of language input. This study takes the findings from Le Normand et al. (2013) one step further by looking at the relationship between SES and content and function words instead of only function words, while still analyzing how function words impact MLU. Current research is reviewed regarding SES, quality and quantity of the input, and how content and function words are impacted in this early developmental period. Additionally, research conducted in museum settings are discussed in relation to early language development and SES.

The role of SES in child language development

On average, children are able to produce about 300 words by age two (Fenson et al., 1994), and reach an expressive vocabulary of close to 14,000 words by age six (Clark, 1995). However, these numbers are affected by SES as children who grow up in poverty are exposed to approximately 30 million fewer words by age three than those who are raised in higher SES families (Hart & Risley, 1995). The lack of expressive language by these children is due to factors including parents' inability to afford to stay-at-home with their children and spend time conversing with them, lower levels of maternal education, and limited access to books and other forms of literature in the home (Debaryshe, 1993).

A lack of literature in the home is especially detrimental to children's language development because it is through the shared action of reading picture books and telling stories that parents engage with their child to participate and learn about new words (Rogoff, 1990). Dwyer, Jones, Davis, Kitamura, and Ching (2019) found that infants whose mother had a higher level of education were exposed to more adult speech input, had the chance to engage in more conversational turns, and as a result produced more vocalizations themselves. The increased input that these children receive expands their vocabulary and language production as they get older, resulting in longer utterances (measured with mean length of their utterance or MLU). This increased level of input allows the child to develop a vocabulary for use in different settings, objects, and with different people. The more vocabulary the child learns and has in their lexicon, the greater ability they have to converse about a wider variety of topics. This translates into allowing for increased learning opportunities (Duncan et al., 2013).

In addition to the amount of input a child receives, the quality of input also varies with SES. Quality input includes a more diverse vocabulary, the use of more complex sentence structures, and the academic use of grammar. Children whose parents have lower levels of

education may not hear as much language nor will they hear as high a quality of language in their home conversations. Additionally, because a child is not exposed to all possible utterances belonging to the community language but only to the small sample brought to them by their family, friends, and television (Brown, 1960), this can impact their ability to academically compete with other students who are more advanced in their language development due to the more extensive input they receive at home.

Input quality goes beyond hearing an extensive vocabulary and complex speech and grammatical structure, as it encompasses what is called "referential transparency" and the ability to learn through linguistic context (Cartmill et al., 2013). Referential transparency refers to the ability of a child to hear a new word and apply that word and its meaning to something either within the context of the conversation or to a physical object or thing that is visible during the conversation. Referential transparency is dependent on how clear the new word and its meaning is made in reference to another object or thing. There are situational contexts in which learning a new word can be easier due to the environment in which it is presented. Therefore, children from high SES families may have an advantage in learning new words due to their exposure to a broader range of situational settings combined with their parents having a higher tendency to converse with their children which will likely boost their ability to learn new words in new contexts. The more often that parents speak to their children, the greater the chance that their child will be able to apply new vocabulary words to different contexts because these words are introduced and used in various settings (e.g., an environment in which a new word for an object can be applied visually) (Cartmill et al., 2013). For example, a child may better understand what a hot air balloon is and does by hearing people talk about it while looking at one either in person or in a picture, versus during a conversation about a hot air balloon where there is no visual

representation present. Therefore, it is imperative that parents use contextual situations to teach their children new words to expand their vocabulary and improve their language use. Furthermore, continuous additions to one's lexicon allows for more complete sentence structure and the ability to produce complex sentences.

SES has also been shown to affect syntactic development. Vasilyeva, Waterfall, and Huttenlocher (2008) conducted a longitudinal study to examine the development of syntactic skills of 45 children over the course of 20 months (from age 22-42 months). The authors asked whether or not there is variability among children from various SES backgrounds during the beginning stages of syntactic development, and whether SES has an impact on the development of complex sentence structures as the child gets older and begins to produce various sentence types. This study takes the syntactic literature one step further as it helps to develop a more systematic picture of the impact of SES on syntax. No prior study has made it clear as to whether different aspects of the syntax develop in relation to differences in SES.

Vasileyeva et al. (2008) analyzed the language of 45 native English speakers (22 female) representing diverse populations from the greater Chicago area. Parents were asked to report on their child's age, sex, as well as information about their income, level of education, and occupation. The children were divided into three SES groups based on their parent's level of education. Group 1 consisted of families where the greatest education level achieved was a high school diploma. Group 2 included families where the parents graduated from a 4-year college, and Group 3's parents had obtained a post-graduate degree. Visits were made every 4 months to each of the families over the course of a 20-month timeframe. The visits consisted of the caregiver and the child being videotaped for 90 minutes while they participated in their daily

activities. These tapes were later transcribed and linguistically coded for their intonation contour, number of clauses (*zero*, *one*, or *more*), and correct grammatical structure.

Vasilyeva et al. (2008) measured the amount of grammatically complete complex sentences produced compared to the number of simple sentences produced by the child in each of the individual sessions. Second, they calculated the overall number of simple sentences in each of the utterances to examine the use of three basic sentence types; declaratives, imperatives, and interrogatives. Vasilyeva et al. (2008) found that at an early age (22 months) none of the groups varied from one another when it came to formulate simple sentences. This similarity across groups can be explained by education levels not impacting how children learn the building blocks of sentences. However, the vast majority of children who began producing complex sentences at 26 months old were from the groups where the parents had a higher level of education. Additionally, the amount of speech produced in each group varied in terms of educational category where the children in Groups 2 and 3 produced more speech than Group 1. When all of the children reached 34-months of age, over 95% of the children were producing grammatically correct complex sentences. Therefore, age has a significant influence on language development as most of the participants were able to produce the same kinds of sentences by the end of the study. SES impacts the rate of development as higher parental education levels pave the way for children to use complex sentences and produce a greater amount of speech at an earlier age.

The environments in which children grow up influence their development in terms of the size of their lexicon, grammatical use, and complex sentence structures (Brown, 1960; Cartmill et al., 2013; Duncan et al., 2013; Rogoff, 1990; Vasilyeva et al., 2008). Further evidence is required to narrow down what specific factors in these environments impact the rate of

development of speech and language. Thus, we will not be able to make the direct assumption that the rate of language development is dependent upon parental education levels until we have more research evidence that looks into details such as how education level influences children's production of specific types of words (content or function words).

Content and Function Words

Content words carry meaning (e.g., nouns, verbs, adverbs, adjectives) and are distinguishable from function words which provide sentence structure. These words embody the connotation of a sentence. In comparison, function words represent the functional and grammatical words that help to form a sentence (e.g., determiners, prepositions, conjunctions). Together, these two general word categories form cohesive, grammatical sentences. As English-speaking children develop their language, they tend to learn words that they can match an object to first (Benedict, 1979). Our culture places emphasis on naming objects when speaking to infants and young children which contributes to the faster mapping they develop for nouns and other content words (Benedict, 1979). Grammar-specific words, or function words, tend to be acquired later because they do not have a visual-spatial representation of their meaning and therefore more time and exposure is required to learn how to use these words properly within language. Additionally, prosody, or the rhythm and melody of language, is another reason as to why function words tend to be learned later in the English language.

Prosody is represented through the pitch, stress, pauses, and lengths of utterances in speech (Hirschberg, 2002). These factors are used in speech to provide emphasis in order to share information, show emotion, and produce different kinds of utterances (e.g., declaratives, questions, etc.). In the English language, stress is often placed on syllables or words to place significance on a segment of a word or sentence (Dowhower, 1991). Often, these stressed

syllables are placed on content words like nouns or verbs (Whalley & Hansen, 2006). Therefore, this prosodic marking highlights information that children need to learn new words. It also helps explain why children tend to learn content words before they acquire function words due to the prosodic emphasis placed on specific syllables in the everyday language that they hear and use (Cutler & Swinney, 1987).

Le Normand, Moreno-Torres, Parisse, and Dellatolas (2013) performed a corpus analysis to look at how grammatical words are learned and used by children, and how SES and age impact the acquisition and use of grammar and MLU. Participants included 312 parent-child dyads with the children ranging in age from 2 to 4 years old (170 male). All participants were monolingual native French speakers and scored in the normal range on a nonverbal cognitive assessment and passed a hearing screening. Family SES was determined using a categorization developed by Desrosières, Goy, and Thévenot (1983) which took into account the family's income, the mother's level of education, and whether the father was employed.

Le Normand et al. (2013) collected 20-minute language sample recordings of parent-child dyads in a naturalistic play setting. The naturalistic play setting was chosen to motivate the child to produce conversational-like speech during their play with a familiar adult. The 20-minute recordings were transcribed using CLAN (Child Language Analysis; MacWhinney, 2000), a program for analyzing transcriptions. The individual word categories (grammatical, pragmatic, lexical) as well as the subsets of types of grammatical words were tagged using the POS-T software, which is an automatic syntactic analysis system. Statistical analyses examined factors that might influence the children's grammatical word use such as SES, age, and MLU (Le Normand et al. 2013).

Le Normand et al. (2013) found MLU growth to be primarily dependent on the amount of grammatical and pragmatic word types produced by a child. The most common grammatical word types were subject-pronouns, determiners, and prepositions. The usage of more grammatical words in children's speech lends hand to a higher MLU compared to lexical and pragmatic word types, and this was true across all age groups. The use of the amount of variation among word types (subject-pronouns, determiners, prepositions) produced across the age groups were significant in relation to one another, however this was not true across SES groups. Therefore, the older the child, the more prevalent grammatical word types were in their speech. Regarding SES, there was a significant relationship in terms of MLU, however, there was no significant relationship between SES and the types of grammatical words produced in speech because within age groups, SES did not demonstrate the ability to increase a child's natural use of different function words. This result suggests that while SES may speed up the rate of language development overall, the process by which language is developed remains the same. Thus, the types of grammatical words learned will still be acquired in the same order of increasing complexity as defined by Brown (1973).

The current study extends Le Normand et al. (2013) by analyzing similar aspects in English in order to test the generalizability of their findings. Since SES can impact the amount of input, it can thus have an indirect impact on the types of words that children produce most frequently (Hart et al., 1995). The current study uses a similar approach of looking at the types of function words (e.g., determiners, prepositions, conjunctions) and how frequently they are used in child speech during naturalistic play with a parent/caregiver. The results are compared with the findings from Le Normand et al. (2013) to assess if there are associations in the factors

assessed and whether there are similarities in the grammatical word usage and MLU across French and English-speaking children.

Family-oriented studies conducted in museums

An educational-oriented museum like the Children's Museum of New Hampshire (CMNH) supports caregiver-child play in meaningful ways that promote learning. CMNH facilitates learning via the use of materials that explore various points in time, different cultures, and animals with their corresponding habitats that are native to the state of New Hampshire. Caregivers can explore these diverse exhibits and learn alongside their children while taking on the role of 'teacher' and intervening when appropriate to educate their child (Hensel, 1987). Project ACME (Advancing Children's Museum Engagement) utilizes two of the CMNH's exhibits, one that is specific to learning about patterns (and the medieval time point in history), and one that focuses on a native river ecosystem and those who inhabit it. These two areas of the museum encourage families to participate in pretend play and conversation in order to learn about patterns, knights, castles, the migration of fish, beavers, etc.

Other studies have found that families who participate in museum studies tend to engage with their children in ways that promote learning about the topics available in the museum. Additionally, observations done on the conversations that take place among family members while touring museums found that the discussions allowed family members to reinforce old experiences, explore family history, and develop shared understandings involving the interaction with an exhibition (Ellenbogen, Luke, & Dierking, 2004). Luke, Coles, and Falk (1998) analyzed families in a museum environment and discovered that the families continued to discuss the things that they had seen and learned about during their visit over the course of the next few weeks. This effort to point out what was learned in the museum to real-life scenarios outside of

the museum helps to develop the child's ability to fast map, or understand a newly learned concept or word by applying it to other areas outside of where the concept/word was learned (Heibeck & Markman, 1987).

The Current Study

Over the course of the last 50 years, the lexical nature of young children's multi-word utterances has been disputed among researchers (Ninio, 2006; Tomasello, 2000). Tomasello (2000) believes that novel uses of grammar is not used until an individual is 4 years old, while Ninio (2006) argues that children begin to facilitate the use of grammar in their speech around 3 years old. This research is still inconclusive as to how SES affects a child's specific lexical productions. With regard to SES, it is well known that it affects the rate of lexical development, but it is not known the extent to which SES impacts the acquisition and use of specific word types (Vasilyeva, Waterfall, & Huttenlocher, 2008). This study will ask if apart from rate of development, are there more qualitative lexical differences among children from different SES groups. Specifically, it will analyze the amount of content words, as well as the amount of function words that children from the New Hampshire/Maine seacoast area aged 3 to 6 years old produce relative to their SES. Additionally, the ratio of each type of function word relative to the total number of function words will be analzyed in relation to SES, MLUm, and age group.

The primary goal of this study is to better understand how children who come from a variety of SES backgrounds use content and function words in their language production. It will contribute to the area of research concerning language development in the English language and report on the types of words that are less common as well as most common in early language development. There are three research questions central to this study: 1) how do children from high and low SES households compare in the types of words (function or content) that they

produce? 2) does the relationship between the use of function words and MLU vary across SES groups? and 3) is there a unique type of grammatical word that is more strongly related to MLU for the English language?

Based on the existing literature regarding SES, the first hypothesis is that children who come from families of higher SES, will have a better knowledge of how to use grammatical language in speech, and thus produce more function words than children from lower SES families. Research shows that the amount of quality input received by a child from a high SES household tends to be broader than what is received on average for children from lower SES backgrounds (Dwyer et al., 2019). Language cannot be learned without exposure, thus children from a higher SES will have more exposure to language in general. The more function words that a child is exposed to, the greater the likelihood that these words will generalize into the child's vocabulary. In general, based on research regarding language production and SES, the more exposure a child has to specific words, the more confident one can be that they will learn to use the new words, leading to the development of a comprehensive and more robust lexicon.

The hypothesis in regard to the second research question is that an increase in grammatical diversity (function word use) should increase corresponding MLU. SES is expected to have an impact on the rate of development of langauge for these children as well. This hypothesis is based on the research that English-speaking children's early vocabularies have a predominance of nouns compared to other words (Benedict, 1979). However, multi-word utterances rely upon grammatical words in order to link the content words to be able to form a meaningful utterance. Therefore, if a child has the ability to use a wider variety of grammatical words in their lexicon, they will be able to produce longer sentences or multiword utterances (Conwell & Demuth, 2007). Results from other studies suggest that language development and

academic achievement in children are positively correlated by their family's SES, therefore children of higher SES backgrounds must be exposed to high-quality learning environments more often than their low-quality peers that allow them to learn and further their language development (Debaryshe, 1994; Hart & Risley, 1995; Walker, Greenwood, Hart & Carta, 1994). Thus, combining these two aspects suggests that if a child has greater access to high quality and high quantity levels of language, they will learn a larger variety of word forms that will generalize to their expressive language use. Specifically looking at function word development, the results can be compared back to Le Normand et al. (2013)'s results to assess whether English-speaking children of similar ages speak in a manner in which their grammatical language increases the average MLUm per SES category.

The third hypothesis is that the most common types of function words found in the participant's speech and language samples should have the strongest correlation to MLU. Le Normand et al. (2013) found pronouns, determiners, and prepositions to be the best predictors in the French language. This suggests that some grammatical word classes are more actively associated with MLU. Additionally, determiners, prepositions and pronouns are some of the most frequently used words in the English language like in French, which attributes to why the use of these types of words are important for development and utterance length (Le Normand et al., 2013). Additionally, function words do not carry any conceptually complicated subject matter. Therefore, if following Naigles (2002)'s theory that "form is easy but meaning is hard", then it follows that the development of function words should come easily to young children who do not necessarily need to understand the meaning of the word they are using to create grammatical relationships within their speech.

To test these hypotheses, the use of content and function words in early language production will be assessed as SES varies. Importantly, this study will examine naturalistic language samples of children aged 3 to 6 years old. The language samples are taken from play sessions between a caregiver/parent and their child in a children's museum. The findings from this study will further the results of Le Normand et al. (2013), which showed that French-speaking children who built their multiword utterances around specific word types (e.g., content or function words) exhibit a variation in MLU based on an increase in such word types. Le Normand et al. (2013) emphasized that more empirical research is needed to understand how children build early multiword utterances, which should ideally be conducted with children from a variety of linguistic backgrounds. The comparison of the results from this present study would help to generalize the effects of SES on word production as a language independent phenomenon. Expanding upon Normand et al. (2013), this work elaborates on how SES specifically impacts English-language speakers' use of word forms to create their multiword utterances in a larger cohort of children.

Methods

Data for this study is taken from Project ACME (Advancing Children's Museum Engagement), a larger study conducted at the Children's Museum of New Hampshire (CMNH) in Dover, NH. Project ACME looks at a large set of variables related to caregiver/child interactions, behavior, and language by observing caregiver/child behaviors in a naturalistic play setting. The project utilizes headset microphones and video cameras to record the caregiver and the child's verbal engagement. Importantly for the current study, Project ACME attracts a variety of families from different SES backgrounds in Dover and the surrounding seacoast communities.

Participants

Participating dyads were recruited through the museum's website and social media pages (e.g., Facebook), flyers posted in local shops and bulletins across the community, and via word of mouth. Criteria for children participating in the study required them to be between the ages of 3 and 6 years old, have no speech/language, genetic, developmental, or neurological disorders, be Native-English speakers, have normal or corrected-to-normal vision/hearing, and able to provide verbal assent. Upon arrival for participation, the researchers introduced themselves and described the itinerary for the visit. Parents/caregivers of the participating dyads were required to sign an informed consent form before the study began as well as complete. Additionally, they were asked to complete family and child demographic forms to the best of their ability/comfort level. The family demographic form asked questions regarding the caregiver's age, marital status, education level, income level, ethnicity, etc. The child demographic form asks fewer questions such as the child's age, ethnicity, and gender identity. Copies of the family and child demographic consent forms are attached (Appendix A). The participating child was then administered the Clinical Evaluation of Language Fundamentals-Preschool, Second Edition (CELF-P2; Wiig, Wayne, & Semel, 2004), an assessment designed to analyze a child's language skills between the ages of 3 and 6;11. This assessment is often used to establish baseline language skills. The children only were required to complete the sentence structure, word structure, and expressive vocabulary sections of the CELF-P2 which are the sub-tests necessary to yield a Core Language score. The participants' SES level was assessed using the classification of maternal education level. Those who consented to participate receive free admission to the Museum for the day that they participate in the study, and upon successful completion of all research tasks (demographic forms and videotaped interactions) they received a $25 gift card as compensation. Participants are free to withdraw from the study at any time and will still receive

free admission to the museum, but not the gift card. The visits typically last between 1-2 hours, so the compensation provided is equal to over $10 an hour.

Thirty parent-child dyads were selected and analyzed from the larger data set of Project ACME. The children's ages ranged between 3 and 6 years old with the average age of 4 years, 11 months. There were 8 males and 22 females who participated with their caregivers, and two socioeconomic strata present (high and low). Participant selection from the ACME data pool was made based on whether both the audio files from the two exhibits (Castle and River) were intelligible, didn't contain static, and ones in which the child wore the headset. If these standards were met, good/great audio recordings were achieved and able to be adequately transcribed. Additionally, we only used participants whose parent completed the questions regarding SES in the family demographic form (Appendix A). Data were excluded if 1) there was poor audio-headset transmission (# excluded), 2) the child refused to wear the audio headset (#), or 3) there were technical difficulties with recording software Garage Band, which was used to record the audio for the study (#). Out of 47 potential dyads, the final analysis for this study included 30 child/caregiver participants.

SES groupings were based on maternal education (ME). High SES was given to families whose mother had obtained an associate degree/technical certification, bachelor's degree, or an advanced degree. Families were identified as low SES if the mother's education level consisted of some college, a high school degree or GED, or if they attended high school but never graduated. Income was not able to be considered a factor in this study because the large majority of the group sample fell into a high-SES category of income level (making more than $100,000). After the participants were placed into their respective groups, there were 21 dyads who fell into the high SES category, and 9 who were considered low SES based on ME.

Commented [TJ1]: May not want to include this at all, but keep for now

Materials

In addition to the structures and toys already present in the museum, Project ACME provided supplementary materials to encourage play between the dyads, as well as to provide novel materials for families who had been to the museum before. There were two exhibits utilized in this study: the Cocheco River exhibit and the Castle exhibit. In the Cocheco River exhibit, close-ended play materials were provided in the form of raft building materials, beaver-backpacks, and a magnetic fishing game. The Castle exhibit included open-ended style toys (e.g., blocks, felt boards, castle figurines). Open-ended toys do not have a defined ending point, rather they can be played with in a multitude of ways which hopefully promotes the use of imagination, and facilitates more language (Liu, 2018). The Castle exhibit was visually recorded using two GoPro cameras set up in opposite corners of the room while the participants wore headset microphones to have an accurate recording of their speech productions. For the purpose of this study, the video recordings were not necessary for use as part of analysis. Samson XPD1 Headset USB Digital Wireless System audio headsets were used to collect the language samples from the dyads. The child and adult were required to each wear their own headset to collect the best possible quality language sample. The headsets came with a wireless receiver built into a USB stick that was plugged into a Mac laptop where Garageband was opened to record and save the audio files collected. For this study, only recordings from the Castle exhibit were analyzed. This selection was made to avoid any discrepancies in expressive language caused by recording language/speech from two different environments. Furthermore, only the child's language samples were studied since it is their speech that the research questions are looking to observe.

Procedure

Previous studies have assessed language samples by identifying a specific period of time in which to analyze speech from the larger sample (Crystal, Fletcher, & Garman, 1976; Tyack, & Gottsleban, 1977). This approach was used for this study with 15-minute language samples selected from the full 20-minute play session from the Castle exhibit. For the 30 participant dyads, there was a total of 7.5 hours of speech sampled. This is considered a reliable sample size for English speakers for the age group analyzed (Gavin & Giles, 1996).

The caregiver-child dyads each participated in two naturalistic play setting environments in the museum during the study. One of the settings was in the "Cocheco River" exhibit in which informational learning materials are provided to help facilitate conversation and play regarding the animals and plants found in New Hampshire's river ecosystems. Part of this exhibit also yielded information on how the Cocheco River was formerly a main source for the deliverance of goods. The second exhibit was the "Castle" exhibit which was approximately a third of the size of the River exhibit and contained learning materials that concerned abstract shapes, different colors, and patternmaking. The caregiver-child dyads were given a total of 20 minutes to play in each exhibit, with this study analyzing the first 15 minutes of play since not every dyad utilized the entire 20 minutes allotted. The majority of play was on the ground, where the dyads could engage eye-to-eye. The exhibits were blocked off during the study to allow the participants uninterrupted play.

Language Sample Analysis

The mean length of utterance as counted by the number of morphemes (the smallest meaningful unit in a language), or MLUm, was utilized since it is a more accurate method of measuring language development in child language acquisition than age (Brown, 1973). It is calculated by dividing the total number of morphemes from all utterances of an entire language

sample by the total number of utterances produced (Klee & Fitzgerald, 1985; Rondal, Ghiotto, Bredart, & Bachelet, 1987). MLU is a valid measure of development especially for the participants in this study who range from 3 to 6 years old, as MLU is most reliable for children 4 years old and younger. Additionally, MLU has an 85% reliability level for slightly older children between the ages of 5 and 6 years old when the language sample used for analysis is of adequate length (i.e., around 50 utterances) (Rondal et al., 1987; Darley & Moll, 1960). Generally, increasing the size of the sample translates to an increased reliability of generalizing the outcomes or impact to a certain population. This is true in the case of language samples as well, as research has shown that at least 50 utterances should be acquired to achieve appropriate reliability scores in terms of mean length of response (e.g., MLR, a similar metric to MLU word or MLUw which measures mean length of utterance in words by dividing the number of words by the number of utterances) for young children (Darley & Moll, 1960; Rondal, & DeFays, 1978). This study's data are collected from language samples that all exceed 50 utterances, with most exceeding 100 (27 of the 30 participants exceeded 100, 4 exceeded 200 utterances), as can be observed in Table 1 below. The average MLUm produced by each age group is evident in Table #2.

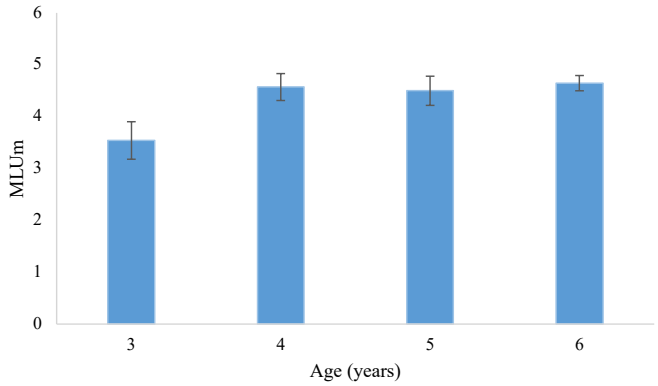
Table 1.

Number of participants and mean MLUm per age group

Age (years)	# of participants	*Mean* MLU (SD)
3	5	3.54 (0.81)
4	12	4.57 (0.90)
5	9	4.5 (0.84)
6	4	4.65 (0.29)
ALL	**30**	**4.32**

Figure 1.

Bar graph of mean MLUm per age group



The language samples were each transcribed in Praat, a computer software program designed for the manipulation, and analysis of speech (Boersma & Weenink, 2018). The analysis of the speech samples was completed using SALT (Systematic Analysis of Language Transcripts) which is intended specifically for transcribing and analyzing language samples (Miller, 1993). To maintain uniformity throughout the transcriptions, SALT's suggested transcription conventions were followed. SALT was used to calculate MLUm, and then calculate the amount of each type of content and function word. The content words categories analyzed were nouns, adjectives, verbs, and adverbs. The function words analyzed were prepositions, auxiliaries, determiners, coordinators, and pronouns. The SALT software creates categories for each word spoken by the child and caregiver during the language sample and tallies the total number of words said by each individual. Four trained research assistants transcribed the recorded language samples in Praat and then transferred the transcriptions to Word documents where they were then uploaded into SALT for analyses. The research assistants compared their transcriptions to one another and came up with 87.5% inter-rater reliability. Any discrepancies were discussed in order to arrive at a consensus for any disagreements.

Statistical Analysis

Correlations and independent sample t-tests were performed using SPSS Statistical software. The independent variables were SES and age. The dependent variables were the number of content words, number of function words, different types of function words, and MLU. The distribution of the 30 participant samples included two SES levels (low: 9 samples, high: 21 samples). Unfortunately the two SES groups are largely variant from one another therefore, to acheive statistical measures that are comparable and reliable 9 high SES participants were selected to match the pre-existing 9 participants from the low SES group. The 9 high SES participants were chosen based on their closeness in age and relative CELF scores (Table 2). The age range of the 18 remaining participants was 3;4-5;9, *M* = 4;11. There were 11 females, 7 males.

Table 2.

Matched Pairs Based on Age and CELF Scores

	Low SES	**High SES**	***t-test***	***p***
Mean Age (SD)	55.33 (9.29)	55.22 (10.33)	0.024	0.981
Mean CELF Score (SD)	106.78 (12.06)	112.89 (10.37)	-1.152	0.266

Results

The first analysis examined the types of words (content or function) that children from high versus low SES backgrounds produce relative to one another. An independent samples t-test discovered no significant difference in the amount of function words produced by the different SES groups, $M = 4.5$, 95%, CI [13.23,4.12], $t(16) = 1.11$, $p = 0.28$. However, an independent samples t-test did find a statistically significant difference in the amount of content words produced by high versus low SES groups, $M= 25.2$, 95% CI [48.22, 2.22], $t(16) = 2.32$, $p = .034$, with the high SES group using more content words than the low SES group.

A Pearson's product-moment correlation was conducted comparing the age of the participants (months) to the number of function and content word types (represented in Figures 2 and 3). Preliminary analyses showed the relationship to be linear with both variables normally distributed, as assessed by Shapiro-Wilk's test (p = .424), and there were no outliers. There was a statistically significant, moderate positive correlation between participant age and function word production, $r(18)$ = .49, p = .039, with differences in age statistically explaining 24% of the variability in the amount of function word types produced. Additionally, a Pearson's product-moment correlation was run to assess the relationship between participant age (months) and the amount of content word types produced by the matched groups. Preliminary analyses showed the relationship to be linear with both variables normally distributed, as assessed by Shapiro-Wilk's test (p = .94) and there were no outliers. There was a moderate positive correlation between age and the number of types of content words produced by the matched SES groups of participants aged 3-6 years, r = .32, however, it was not statistically significant, p = .193.

Table 3.

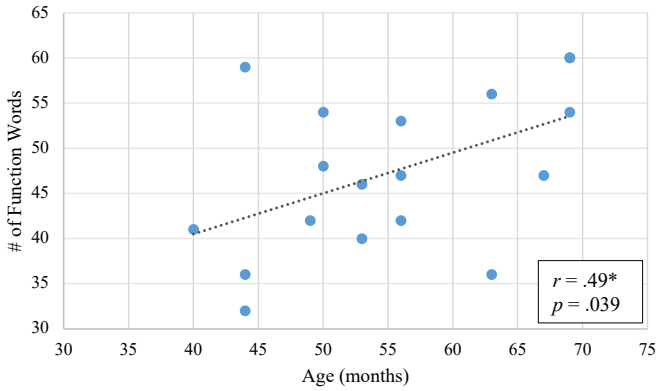
Pearson Correlations for Age and Content/Function Words

	Age (months)
Content words	0.321
Function words	0.491*

* = $p < .05$

Figure 2.

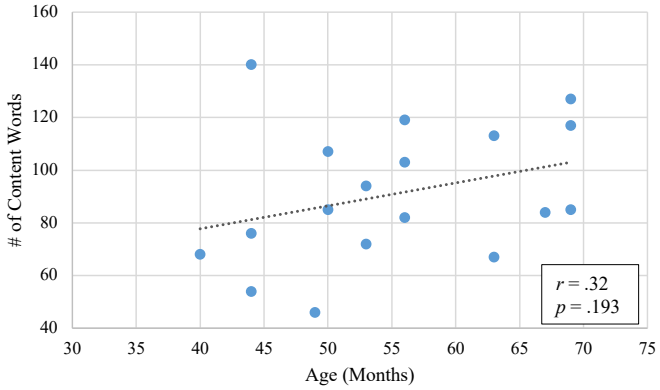
Correlation Between Age and Function Word Use



Note. * Correlation is significant at the 0.05 level.

Figure 3

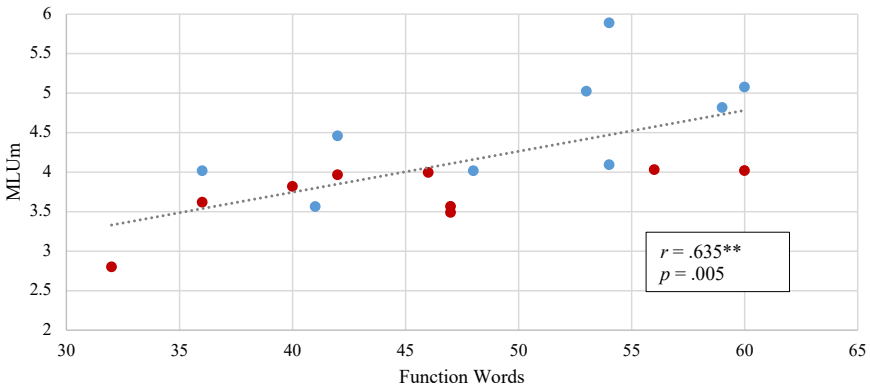
Correlation Between Age and Content Word Use.



The relationship between function word use and MLU is represented in Figure 4. A Welch's t-test was run to compare MLU and its relation to SES. Results showed that $p = .009$, rejecting the null hypothesis and there is a statistically significant difference between MLU and SES.

Figure 4.

Relationship Between Function Word use and MLU Corresponding to SES



Note. Red dots represent low SES participants, blue dots represent high SES participants. ** Correlation is significant at the 0.01 level.

A correlational analysis was performed for the five grammatical categories and MLUm (Table 4). Preliminary analyses showed the relationship to be linear for all variables except coordinators were not normally distributed as assessed by Shapiro-Wilk's test (p = .000), and there were no outliers. However, a pearson's correlation was still run because the test is somewhat robust to deviations from normality. There was a statistically significant, moderate positive correlation between MLU and determiners, $r(18)$ = .65, p = .003, prepositions $r(18)$ = .53, p = .025, and pronouns $r(18)$ = .49, p = .038. As shown in Table 4, the highest correlations were observed for determiners *(r* = 0.65*),* prepositions *(r* = 0.53*),* and pronouns *(r* = 0.49*)*. These correlations showed that determiners accounted for 42% of the variance in MLU, prepositions for 28% and pronouns for 24%.

Table 4.

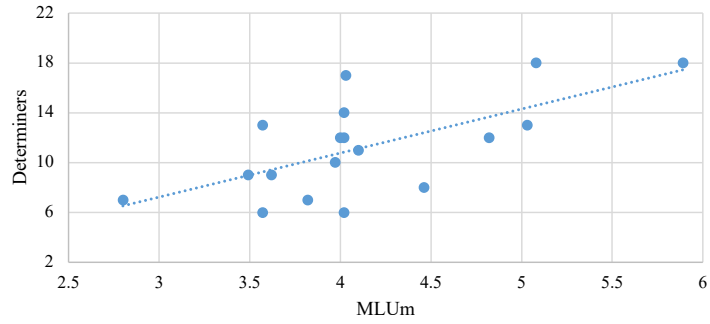
Correlational Analysis of 5 Grammatical Categories

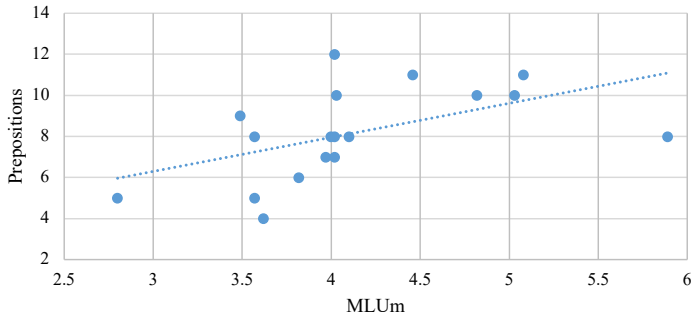
	Mean	SD	Range	r
MLUm	4.16	0.73	2.8- 5.89	
Determiners	11.22	3.86	6.0- 18.0	**0.65**
Prepositions	8.17	2.56	4.0- 12.0	**0.53**
Auxiliaries	10	2.99	6.0- 17.0	0.23
Coordinators	2.1	0.73	0.0- 3.0	0.34
Pronouns	15.78	1.83	13.0- 19.0	**0.49**

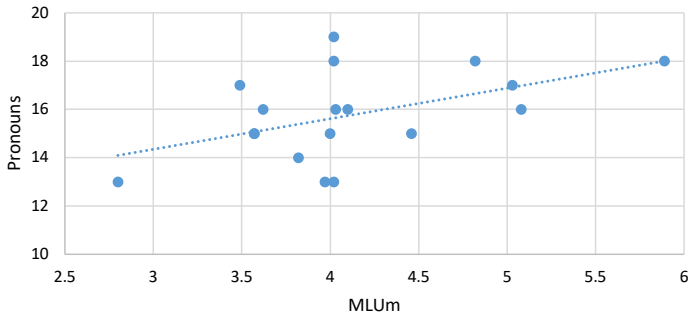
Note. Bold indicates significance $p < .05$

Figure 3.

Correlations Between MLU and 3 Most Common Function Word Types







Note: Correlations between MLUm and number of different determiners (top), prepositions (center), and pronouns (bottom).

To examine the impact age and SES on the individual grammatical categories, correlations between the independent variables were examined: two of the three most identified grammatical categories were positively correlated with age (determiners $r = .57$, prepositions $r = .64$, respectively, $p < .05$), but not with SES (determiners $r = .09$, $p = .73$, prepositions r = .23, $p = .36$, pronouns r = .25, $p = .32$.

Discussion

The main aim of this study was to examine the language development of children in the seacoast region of Maine and New Hampshire in terms of how SES impacts specific word use, and vocabulary development. Three hypotheses have been explored in regards to this area of study. To assess these three hypotheses, a group of 30 children ages 3-6 were analyzed using language samples that range from 60-350 utterances per child. In order to account for the different number of participants in each SES group, two groups of 9 children each (9 high SES, 9 low SES) were matched on age and CELF scores.

According to the first hypothesis, children of higher SES will produce function words based on the idea that they would have better knowledge of grammatical language use. An

important finding was the types of words that are more prevalent between SES groups. Content words significantly correlated with SES while function words did not significantly correlate with SES. Thus, SES does not seem to impact the amount of grammatical words. Interestingly, children of higher SES are using more content words in their speech. The hypothesis that children of high SES would produce a greater amount of function words than children coming from families of a lower SES was not upheld. However, this finding aligns with results from other studies where the quality of input and quantity of input, both of which tend to be greater in higher SES families, have a positive influence on children's ability to learn new words and expand their lexicon (Duncan et al., 2013; Hart & Risley, 1995).

It is important to note that age can still, and should, play a factor in how many content/function words a child produces in their language. It was necessary to determine if age correlated with the number of content and function words produced by each of the 18 children. Analyses showed that the number of function words and the number of content words both showed moderate positive correlations with age. These results suggest that age is most strongly correlated with expressive vocabulary (the number of function and content words) and that older children tend to have larger vocabularies and have developed the use of a wider variety of words. Similarly, Le Normand et al. (2013)'s also found relationships between age and word types to be significant, suggesting that word types vary across age, but not necessarily for SES.

The second hypothesis in this study looked at the relationship between function word use and its impact on MLU in relation to SES (high and low). There is no significant relationship between function words and SES groups, therefore function word use and MLU cannot be compared in terms of how they are impacted by SES. However, MLU measured alone with SES demonstrates a significant relationship. A higher SES corresponds with a higher MLU score. Le

Normand et al. (2013) found a positive relationship between SES and MLU as well, consequently these results support their findings.

According to the third hypothesis, the most common type of function word will be the best predictor of MLU. Three grammatical word types (determiners, prepositions, and pronouns) were the greatest predictors of MLU in this study. This confirmed the hypothesis that certain grammatical word classes are more strongly associated with MLU. These results relate to Le Normand et al. (2013) which also showed that pronouns, determiners, and prepositions are the best predictors of MLU in the French language. In this study, in English, it appears that determiners are the most predictive compared to French in which pronouns are the most predictive of MLU. It is possible that determiners are so predictive because they tend to precede nouns/noun phrases. Within the English language, nouns (object-naming words) are learned first (Benedict, 1979). Therefore, when producing grammatically correct statements and sentences, nouns will be preceded by an appropriate determiner, which might explain why determiners were shown to be the most commonly used function word. Compared with Le Normand et al. (2013) finding of pronouns as the most common function word in the French language, this finding of determiners as the most commonly used function word in the English language further suggests that these findings are of language-dependent nature.

Summary and Conclusion

This study aimed to better understand how children who come from different SES backgrounds use content and function words in their language production, and how function word use impacts MLU. Using t-tests and correlations, this study has explored the relationship SES has with function and content word production and has found determiners to be the most commonly used function word in this sample of American English speakers. The results give

value to the study as a whole and further the results of Le Normand et al. (2013) in the sense that we now understand function words to be used at different levels within languages. The strengths of this study include a sizable sample size (n= 30), and that all child participants are native English speakers, which allows the results to be compared across languages. Limitations of this study include a limited SES diversity, an unequal amount of age groups represented in the participant population, and is not an accurate representation of English language speakers across the United States. Looking at the broad impact, this study informs our understanding of how SES shapes language, highlighting the critical need for these considerations in research and clinical settings.

References

Benedict, H. (1979). Early lexical development: Comprehension and production. *Journal of Child Language, 6*(2), 183-200.

Boersma, P. & Weenink, D. (2018). *Praat: doing phonetics by computer* [Computer program]. Version 6.0.42, retrieved from http://www.praat.org/

Brown, R. W. (1973). *A first language: the early stages.* Cambridge, MA: Harvard University Press.

Brown, R. W., & Berko, J. (1960). Word association and the acquisition of grammar. *Child Development, 31*, 1-14.

Cartmill, E. A., Armstrong, B. F., III, Gleitman, L. R., Goldin-Meadow, S., Medina, T. N., & Trueswell, J. C. (2013). Quality of early parent input predicts child vocabulary 3 years later. *Proceedings of the National Academy of Sciences of the United States of America, 110*(28), 11278.

Clark, E. (1995). Later lexical development and word formation. In P. Fletcher, & B. MacWhinney (Eds.), *The Handbook of Child Language* (pp. 293-312). Oxford: Basil Blackwell.

Crystal, D., Fletcher, P., & Garman, M. (1976). The grammatical analysis of language disability: A procedure for assessment and remediation. *NewYork: Elsevier.*

Conwell, E., & Demuth, K. (2007). Early syntactic productivity: Evidence from dative shift. *Cognition, 103,* 163-179.

Darley, F. L., & Moll, K. L., (1960). Reliability of language measures and sizes of language sample. *Journal of Speech & Hearing Research, 3,* 166-173.

Debaryshe, B. D. (1993). Joint picture-book reading correlates of early oral language skill.

Journal of Child Language, 20(2), 455-61.

Desrosières, A., Goy, A., & Thévenot, L. (1983). L'identité sociale dans le travail statistique: la nouvelle nomencla- ture des professions et catégories socioprofessionnelles. *Economie et Statistique, 152*, 55–81.

Dowhower, S. L. (1991). Speaking of prosody: Fluency's unattended bedfellow. *Theory Into Practice, 30*(3), 165-175).

Duncan, G. J., & Magnuson, K. A. (2003). Off with Hollingshead: Socioeconomic resources, parenting, and child development. *Socioeconomic status, parenting, and child development,* 83-106.

Dwyer, A., Jones, C., Kitamura, C., & Ching, T. (2019). Maternal education influences Australian infants' language experience from six months. *Infancy, 24*(1), 90-100.

Ellenbogen, K. M., Luke, J. J., & Dierking, L. D. (2004). Family learning research in museums: An emerging disciplinary matrix? *Science education, 88*(1), 48-58.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5), 1-185.

Gavin, W., & Giles, L. (1996). Sample size effects on temporal reliability of language sample measures of preschool children. *Journal of Speech, Language, and Hearing Research, 39*(6), 1258-1262.

Hackman, D. A., & Farah, M. J. (2009). Socioeconomic status and the developing brain. *Trends in cognitive sciences, 13*(2), 65-73.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore, MD: Paul H. Brookes Publishing Company.

Heibeck, T., & Markman, E. (1987). Word learning in children: An examination of fast mapping. *Child Development, 58*(4), 1021-1043.

Hensel, K. A. (1987). Families in a museum: Interactions and conversations at displays. *Dissertation Abstracts International, 49*(09). (University Microfilms No. 8824441.)

Hirschberg, J. (2002). Communication and prosody: Functional aspects of prosody. *Speech Communication, 36*(1-2), 31-43.

Klee, T., & Fitzgerald, M. D. (1985). The relation between grammatical development and mean length of utterance in morphemes. *Journal of Child Language, 12*(2), 251-269.

Le Normand, M. T, Moreno-Torres, I., Parisse, C., & Dellatolas, G. (2013). How do children acquire early grammar and build multiword utterances? A corpus study of French children aged 2 to 4. *Child Development, 84*(2), 647-661.

Liu, F. (2018). Application of modularity principles in designing of functional regenerating open-ended toys alongside with child development.

Luke, J., Coles, U., & Falk, J. (1998). *Summative evaluation of DNA zone, St. Louis Science Center. Unpublished evaluation report. Annapolis, MD: Institute for Learning Innovation.*

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk (*Vol. I, 3rd ed.). Mahwah, NJ: Erlbaum.

Miller, J. (1993). *Systematic analysis of language transcripts (SALT 1)* [Computer software, IBM/PC version 3.0]. Madison, WI: University of Wisconsin, Language Analysis Lab.

Mira, W., & Schwanenflugel, P. (2013). The impact of reading expressiveness of the listening comprehension of storybooks by prekindergarten children. *Language, Speech & Hearing Services in Schools (Online), 44*(2), 183-194.

Naigles, L. R. (2002). Form is easy, meaning is hard: Resolving a paradox in early child language. *Cognition, 86,* 157-199.

Ninio, A. (2006). *Language and the learning curve: A new theory of syntactic development.* Oxford, England: Oxford University Press.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context.* New York, NY: Oxford University Press.

Rondal, J. A., & DeFays, D. (1978). Reliability and mean length of utterance as a function of sample size in early language development. *Journal of Genetic Psychology, 133*(2), 305-06.

Rondal, J. A., Ghiotto, M., Bredart, S., & Bachelet, J. (1987). Age-relation, reliability, and grammatical validity of measures of utterance length. *Journal of Child Language, 14*, 433-446.

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition, 74,* 209-253.

Tyack, D., & Gottsleban, R. (1977). Language sampling, analysis and training. *Palo Alto, CA: Consulting Psychologists Press.*

Vasilyeva, M., Waterfall, H., & Huttenlocher, J. (2008). Emergence of syntax: commonalities and differences across children. *Developmental Science,* (1), 84.

Walker, D., Greenwood, C., Hart, B., & Carta, J. (1994). Prediction of School Outcomes Based on Early Language Production and Socioeconomic Factors. *Child Development, 65*(2), 606-621.

Whalley, K., & Hansen, J. (2006). The role of prosodic sensitivity in children's reading development. *Journal of Research in Reading, 29*(3), 288-303.

Wiig, E. H., Wayne, A. S., & Semel, E. (2004). *Clinical Evaluation of Language Fundamentals-Preschool, Second Edition (CELF-P2).*

Appendix A- Family and Child Demographic Information Forms

Family Demographic Information

We would like to ask some questions about you, your family, and home.

What is your relation to the child in the study?

- ○ Mother
- ○ Father
- ○ Legal Guardian (Please specify your relationship to your child _____)

What is your current marital status?

- ○ Married
- ○ Not married, but living with significant other
- ○ Divorced or separated
- ○ Single, never married
- ○ Widowed

What is your current age? _____

What is your ethnicity/race (select all that apply):

- [ ] American Indian or Alaska Native
- [ ] Asian or Asian American
- [ ] Black or African American
- [ ] White or Caucasian
- [ ] Hispanic or Latino
- [ ] Other _____

What is your highest level of education?

- ○ Less than high school
- ○ Some high school
- ○ High school diploma/GED
- ○ Some college
- ○ Associate's degree or technical certificate
- ○ Bachelor's degree
- ○ Some graduate school
- ○ Advanced graduate degree (ex., MA, Ph.D, M.D.)

Are you currently employed?

- ○ Yes; currently employed full-time
- ○ Yes; currently employed part-time
- ○ No

If you are married or living with your significant other, what is their current age? _____

If you are married or living with your significant other, what is their ethnicity/race (select all that apply):

- [ ] American Indian or Alaska Native
- [ ] Asian or Asian American
- [ ] Black or African American
- [ ] White or Caucasian
- [ ] Hispanic or Latino
- [ ] Other _____

If you are married or living with your significant other, what is their highest level of education?

- ○ Less than high school
- ○ Some high school
- ○ High school diploma/GED
- ○ Some college
- ○ Associate's degree or technical certificate
- ○ Bachelor's degree
- ○ Some graduate school
- ○ Advanced graduate degree (ex., MA, Ph.D, M.D.)

If you are married or living with your significant other, are they currently employed?

- ○ Yes; currently employed full-time
- ○ Yes; currently employed part-time
- ○ No

What is your approximate yearly family/household income?_____

Including you and your child, how many people and children currently reside in your home?_____

Is the primary language spoken in your home English?

- ○ Yes
- ○ No (please indicate the primary language spoken in your home _____)

Child Demographic Information

First, we would like to ask you some questions about your child.

Child's gender identity: _____

Child's ethnicity/race (select all that apply):

- ☐ American Indian or Alaska Native
- ☐ Asian or Asian American
- ☐ Black or African American
- ☐ White or Caucasian
- ☐ Hispanic or Latino
- ☐ Other _____

Child's age: _____

How often do you come to the Children's Museum? _____

When you come to the Children's Museum, what do you hope your child will get out of the experience? _____